

⑯ 日本国特許庁 (JP) ⑯ 特許出願公開  
 ⑰ 公開特許公報 (A) 昭57-102699

⑯ Int. Cl.<sup>3</sup>  
 G 10 L 1/00

識別記号 庁内整理番号  
 7350-5D

⑯ 公開 昭和57年(1982)6月25日

発明の数 1  
 審査請求 未請求

(全 6 頁)

⑯ 音声認識装置

⑯ 特 願 昭55-179945  
 ⑯ 出 願 昭55(1980)12月18日  
 ⑯ 発明者 大賀英文  
 門真市大字門真1006番地松下電  
 器産業株式会社内

⑯ 発明者 蔡内秀和  
 門真市大字門真1006番地松下電  
 器産業株式会社内  
 ⑯ 出願人 松下電器産業株式会社  
 門真市大字門真1006番地  
 ⑯ 代理人 弁理士 大島一公

明細書

1. 発明の名称  
 音声認識装置

2. 特許請求の範囲

(1) 音声登録パタンエリアと音声入力パタンエリアとを設け、音声登録パタンエリアにはあらかじめ認識すべき音声単語に対応する音声パタンを格納しておいて、前記登録パタンエリア内の音声パタンと、前記音声入力パタンエリア内の音声パタンとのパタンマッチングを行なうことによつて入力パタンエリア内の音声パタンに対応する音声単語を認識する音声認識装置において、登録すべき音声は2回以上発声することにし、前記入力パタンエリアに新たなパタンが入力される毎に、以前の入力パタンエリア内のパタンを退避する登録パタンエリア(I)を設け、入力パタンエリア内の音声パタンと登録パタンエリア(I)内の音声パタンとのパタンマッチングを行なつて類似度を算出し、その類似度が予め定められたシキイ値より大きい時にのみ前記登録パタンエリア(I)内の音声パタンあ

るいは前記入力パタンエリア内の音声パタンを登録パタンエリア(I)へ転送することを特徴とする音声認識装置。

(2) 入力パタンエリア内の音声パタンと以前の入力パタンエリア内のパタンを退避する登録パタンエリア内の音声パタンとのパタンマッチングを行ない、この時の類似度をS<sub>1</sub>とし、さらに前記入力パタンエリア内の音声パタンと、前記登録パタンエリア内の音声パタンとのパタンマッチングを行ない、最大の類似度を検出しこれをS<sub>2</sub>として、前記S<sub>1</sub>がシキイ値より大でかつS<sub>2</sub>がS<sub>1</sub>より大である時のみ、前記パタンを退避する登録パタンエリア内のパタンを登録パタンとして登録パタンエリアへ転送することにした特許請求の範囲第1項記載の音声認識装置。

3. 発明の詳細な説明

本発明は、あらかじめ登録された登録パタン群と入力パタン間とのパタンマッチングによつて、入力パタンが登録パタン群のどのカテゴリに最も似ているかを判断し、入力パタンを識別するパ

ンマッチング識別方法を音声に利用した音声認識装置に関するもので、その目的とするところは登録パタンを登録する場合に正当でないものを排除するよう制御を加えようとするものである。

パタンマッチング方式における音声認識方法について第4図に従つて説明する。発声は孤立音声単語であり、音声はマイク(1)によつて電気信号に変換され、特徴抽出部(2)によつて特徴が抽出されて特徴パタンに変換される。又特徴抽出部(2)では音声区間の検出がなされ、発声された孤立音声単語に対応した特徴パタンは入力パタンエリア(3)に格納される。(5)は認識すべき音声単語の特徴パタンとそのカテゴリコードを格納している登録パタンエリアであり、スイッチ(4)がE(ENTRY)側と接続している場合には入力パタンエリア(3)に格納されている特徴パタンは入力端子(8)より入力されているカテゴリコードとともに登録パタンエリア(5)へ転送される。あらかじめ、スイッチ(4)をE側にして、認識すべき音声単語を発声して、その特徴パタンと、カテゴリコードを登録エリアへ格納す

(3)

なされてリセットされ、リセットに対応したコードが出力される。以上の処理は判定部(7)で行なわれる。

このような音声認識装置においては、登録に関して正当でない登録パタンがそのまま登録されてしまうという問題があつた。登録すべき音声単語を発声した時に、たまたま雑音が加わつて来た場合、あるいはノドの調子がおかしい時に発声した場合等に得られた登録パタンはかならずしも正当なものではない。正当でない登録パタンのままで認識させれば、結果として誤認識が多発することとなる。

本発明は、登録されるパタンが正当なものか否かを判断し、正当なものである場合には登録し、正当なものでない場合には登録しないようにすることにより、従来より問題であつた正当でない登録パタンによる誤認識を減少させたものである。

以下、実施例として示した図面に従つて、その構成を説明する。

(1)のマイク、(2)の特徴抽出部、(3)の入力パタン

る。(4)は入力パタンエリア(3)の入力パタンと登録パタンエリア(5)内の複数の登録パタンとのパタンマッチングを行ない、登録パタン毎に入力パタンとの類似度を出力線(6)を介して出力するパタンマッチング部である。この時には、スイッチ(4)はR(Recognition)側と接続される。(7)はパタンマッチング部(6)からの出力である各登録パタン毎の類似度を出力線(9)で受けるとともにその時のカテゴリコードを出力線(8)より受けて、入力パタンがどの登録パタンと最も似ているかを判別する判定部である。登録パタン毎の類似度の中から最も大きい(最も類似性のある)値  $S_{max}$  を選出する。この  $S_{max}$  があらかじめ定められたシキイ値  $K$  より大なる時、

つまり、 $S_{max} > K$  ..... (1) 式

の時に入力パタンは  $S_{max}$  を出力した登録パタンであると識別されて、その登録パタンのカテゴリコードが出力(9)される。 $S_{max}$  が(1)式を満足しない時は登録パタンエリア内に格納されている登録パタンとは異なつた入力パタンが入力されたとみ

(4)

エリア、(5)の登録パタンエリアおよび(6)のパタンマッチング部について第4図と同様な構成で、同じ動作である。

登録時には、同じ音声単語を2回以上発声することとし、カウンタ(8)は何回同一のカテゴリコードを有する音声単語が発声されたかをカウントするものである。特徴抽出部において、音声区間の検出はなされるが、音声単語の終端が検出されると出力線(7)を介してクロックを出力し、カウンタ(8)はそのクロックによつてカウントアップする。すなわち、音声単語を発声するたびにカウントアップされる。一方カウンタ(8)はリセット時検出部(9)からの信号によつてリセットされる。リセット時検出部(9)は入力端子(8)より入力されるカテゴリコード(登録されるべき音声単語に対応している)を受けて、そのカテゴリコードが変化したか否かを検出し、カテゴリコードが変化した場合にはカウンタ(8)をリセットする。同じ音声単語であるなら当然そのカテゴリコードも同一であるため、このような構成によつて登録する音声単語が何回発

(5)

—754—

(6)

声されたかを知ることが出来る。

カウンタ $\theta$ の出力 $\theta$ はスイッチ $\theta$ を制御し、登録すべき音声単語の発声が第1回目であるなら、スイッチ $\theta$ をオフし、2回目以降ならスイッチ $\theta$ はオンさせるように制御する。入力端子(8)のカテゴリコードを変化させない限りスイッチ $\theta$ はオン状態となるため、認識時にカテゴリコードを変化させないようにすればスイッチ $\theta$ はオン状態のままになる。認識時にはスイッチ $\theta$ をオン状態にするように出力線 $\theta$ を制御しても良い。

登録パタンエリア(1)の入力は、入力パタンエリア(3)の出力と接続され、特徴抽出部(2)からの音声パタンが入力パタンエリア(3)へ格納される時に、入力パタンエリア(3)に格納されていた以前のパタンは登録パタンエリア(1)へ転送される。従つて、登録パタンエリア(1)の内容は、1つ前に発声された音声単語の特徴パタンを格納していることとなる。

スイッチ $\theta$ は登録時にはE側に、認識時にはR側に接続される。スイッチ $\theta$ は信号線 $\theta$ をスイッ

(7)

(9)によつて話者に登録されたか否かを出力する。登録されない場合には、さらに話者は同一の音声単語を発音する。この時には3回目の発声となるが2回目の発声に対応したパタンと3回目の発声に対応したパタンとの類似度が同様算出され、判定部(7)で判定されて、前述したような処理がなされる。このようにすることによつて、類似度がシキイ値 $K$ より大となるまで、何回でも同一の音声を発声することとなる。シキイ値 $K$ より大になつた時にはじめて登録パタンエリア(1)(5)には登録パタンが格納されることとなり、結局登録パタンエリアには正当なパタンが格納され、正当でないパタン、例えば周囲雑音が混在したようなパタンは登録されることになる。

認識時には、スイッチ $\theta$ をR側に接続しておき、かつスイッチ $\theta$ をオフしておいて、スイッチ $\theta$ が常にオフするようにしておけば良い。登録パタンエリア(1)(5)と入力パタンエリア(3)とのパタンマッチングをパタンマッチング部(6)で行なつて、判定部(7)では前述したと同様な処理を行なえば、入力

チ $\theta$ へ伝達するか否かを制御するもので、登録時には接続している。このような構成にしておけば登録時に音声単語を2回発生すれば、1回目の音声は登録パタンエリア(1)へ格納され、2回目の音声は入力パタンエリア(3)へ格納されることとなり、パタンマッチング部(6)は第1回目の音声と、第2回目の音声との類似度を算出することができる。判定部(7)は、第4図の判定部と同様な処理を行なうが、この時には登録パタンエリア(1)のみとしか入力パタンとの類似度は算出されないため、前記した $S_{max}$ は、登録パタンエリア(1)と入力パタンとの類似度となる。

(1)式を満足しておれば、判定部(7)は出力線 $\theta$ を介してスイッチ $\theta$ をオンさせ、入力端子(8)のカテゴリコードとともに、登録パタンエリア(1)の内容を登録パタンエリア(1)(5)へ格納する。登録パタンエリア(1)(5)は、第4図の登録パタンエリアと同様認識されるべき音声単語を格納しているエリアである。もし(1)式を満足していなければ出力線 $\theta$ によつてスイッチ $\theta$ をオンさせる。又同時に出力

(8)

パタンの識別が可能となる。 $\theta$ は登録パタンエリア(1)からの出力線である。登録パタンエリア(1)(5)には正当なパタンしか格納されてないため、従来に比べさらに認識率は良くなり、誤認識も減少することになつた。

第2図は、判定部をさらに改良したものである。動作としては判定部 $\theta$ と判定部(7)は同様であるが、認識時と登録時でシキイ値 $K$ を変化させる。スイッチ $\theta$ は認識時にはR側に、登録時にはE側に接続され、認識時にはシキイ値は $K_R$ で登録時には $K_E$ になる。

$$K_E > K_R$$

としておくことにより、登録時にはさらにきびしく登録パタンが正当か、否か判別されることになる。登録パタンエリアには、より正当性のあるパタンしか登録されないことになる。このようにすればさらに認識率の向上、誤認識の減少が可能になる。

又、第8図に示すような構成も考えられる。入力パタンマッチング部(6)は、入力パタンエリア(3)

(9)

の内容と登録パターンエリア(II)(5)の内容とは、認識時登録時に関係なくパターンマッチングを行なうようにし、登録パターンエリア(II)の内容とは、登録時のみ入力パターンとマッチングを計算するようとする。スイッチ判別は従つて登録時のみオンし、認識時にはオフしている。さらに判定部(7)へは登録時のカテゴリコードを示している信号線の内容を登録時のみ加わる様な構成にしておく。スイッチ判別は、登録時のみオンするようにしておく。カテゴリコード判定部は出力線の内容と信号線の内容を比較し、等しい時のみスイッチ判別をオンするように制御線を介してスイッチ判別を制御する。

本発明はこのような構成をとつたので、登録時にはこれから登録しようとするパターンの他にすでに登録されているパターンともふくめて、パターンマッチングを行なうことができ、従つて判定部(7)で選出される最も類似度の高い値  $S_{max}$  は、すでに登録されているパターンにおける類似度もふくめた値になり、この  $S_{max}$  を出力したカテゴリコードが出力線に出力される。このコードがこれから

(1)

又、発声者は出力線の内容から登録されたか否かを判断することができる。指定したカテゴリコードが出力線に出力された場合には、登録されたことになり、逆に出力線のコードと指定したコードとが異なる場合には登録されないことを示す。出力線の内容が信号線の内容と異なる場合には登録しようとする音声単語がすでに登録されている音声単語のどれと近いかを知ることができる。従つて、登録しようとする音声単語を変更する等の処理をすれば良い。例えばイチ(1)とシチ(7)ならシチをナナと発音するように変更する。

さらに判定部の処理を次に示すようにすれば、上記のような似かよつたパターン同士の登録を防ぐことができる。パターンマッチング部から出力される類似度の中から2番目に高い類似度を  $S_{max2}$  とすると

$$S_{max} - S_{max2} > L \quad \dots \dots \quad (2) \text{式}$$

という条件を加えれば良い。このようにすれば、すでに登録されているパターンとすくなくとも類似度が  $L$  以上のものしか登録されないことになる。

登録しようとするカテゴリコードであつた場合にはカテゴリコード判定部はスイッチ判別をオンし、登録パターンエリア(II)の内容を登録パターンエリア(II)(5)へ転送するように出力線に出力する。 $S_{max}$  を出力したカテゴリコードがこれから登録しようとするカテゴリコードでなかつた場合にはカテゴリコード判定部はスイッチ判別をオフするように出力線に出力する。

又、最も高い類似度  $S_{max}$  が  $K$  より小の場合は、前述したと同様リジエクトに対応したコードが出力線に出力され、当然これから登録しようとするカテゴリコードとは異なるためスイッチ判別のオンはなされない。これから登録しようとするパターンが、すでに登録されているパターンの中に良く似ているパターンである場合(例えば、オカヤマとワカヤマ、イチ(1)とシチ(7)、ナハとナラ)にはすでに登録されているパターンの類似度の方がなんらかの原因(発音の不安定さ、あるいは周囲のノイズ)で、高いことがあります。第3図のようにしておけばこのような場合には登録されないことになる。

(2)

従つて誤認識はさらに改良される。

又、第1図、第3図においては登録パターンエリア(II)を設けたがこれを設けず、登録パターンエリアの内容の一部を消去する機能を設けても良い。登録時には登録エリアへ常に転送するようにしておき、登録時の判定の結果、登録してはいけない場合と判明した時に限つて、登録カテゴリコードに対応した登録パターンの内容を消去するようにすれば第1図あるいは第3図と同様な処理が可能となる。又、第1図において、登録時に正当なパターンであると判断された時には、登録パターンエリア(II)内のパターンを登録パターンエリア(II)(5)へ転送せず、入力パターンエリア(3)の内容を登録パターンエリア(II)(5)へ転送するようにしても良い。

本発明は上記のように雑音、あるいは発声の不安定さ等による原因で生じた正当でないパターンは登録されないために、誤認識を減少することができ、登録時に制限を加えることにより、似かよつたパターンの登録を防ぐことができ、安定した認識ができる効果がある。又、判定部のシキイ値を登

03

04

録時には認識時よりきびしくする方向にすれば、より正当なパターンしか登録されないため、誤認識がさらに減少する等の効果が得られる。

#### 4. 図面の簡単な説明

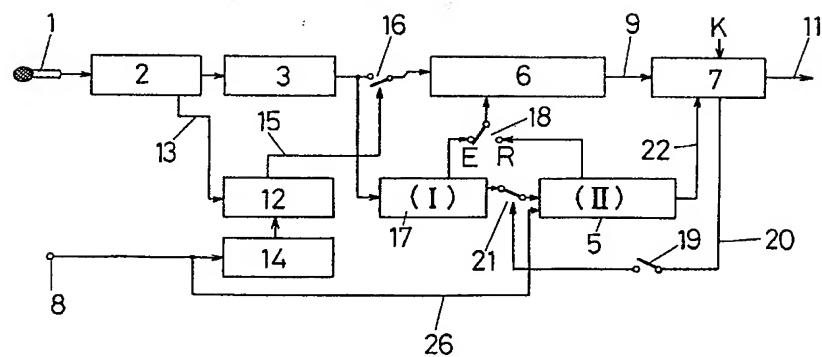
第1図は本発明装置の構成図、第2図は判定部の他実施例を示す構成図、第3図は本発明装置の他実施例を示す構成図、第4図はパタンマッチング方法による音声認識装置の構成図。

(3) … 入力バタンエリア、(5) … 登録バタンエリア  
(1)、(6) … バタンマッチング部、(7) 例 … 判定部、(8)  
… 登録バタンエリア (1)

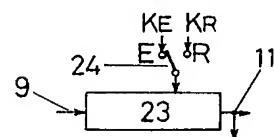
代理人 弁理士 大島 一公

06

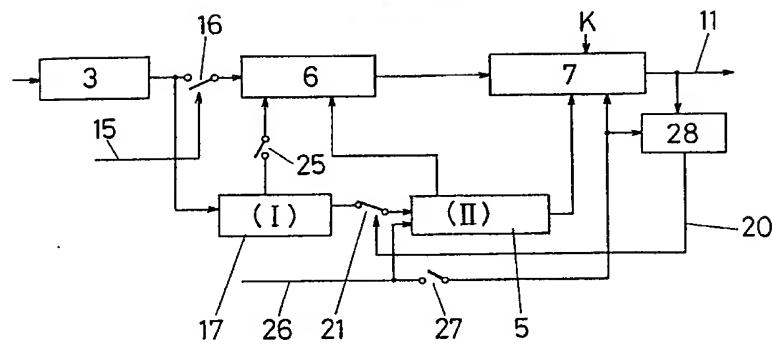
### 第 1 図



## 第2図



第3図



第4図

